

Dynamic decision-making in uncertain environments I. The principle of dynamic utility

Jin Yoshimura · Hiromu Ito · Donald G. Miller III ·
Kei-ichi Tainaka

Received: 3 October 2012 / Accepted: 8 January 2013 / Published online: 26 January 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Understanding the dynamics or sequences of animal behavior usually involves the application of either dynamic programming or stochastic control methodologies. A difficulty of dynamic programming lies in interpreting numerical output, whereas even relatively simple models of stochastic control are notoriously difficult to solve. Here we develop the theory of dynamic decision-making under probabilistic conditions and risks, assuming individual growth rates of body size are expressed as a simple stochastic process. From our analyses we then derive the optimization of dynamic utility, in which the utility of weight gain, given the current body size, is a logarithmic function: hence the fitness function of an individual varies depending on its current body size. The dynamic utility function also shows that animals are universally sensitive to risk and display risk-averse behaviors. Our result proves the traditional use of expected utility theory and game theory in behavioral studies is valid only as a static model.

Keywords Dynamic decision-making · Stochastic environment · Foraging behavior · Risk sensitivity · Expected utility theory

Introduction

Environmental uncertainty and risk are critical factors modulating behavioral sequences in animals. Extensive empirical studies of foraging behavior demonstrate animal behavior is highly risk-sensitive (Real and Caraco 1986; Stephens and Krebs 1986; Stephens et al. 2007; Houston et al. 2011). Mathematical approaches used frequently in traditional studies of foraging behavior include (1) the mean-variance or variance-discount method (Stephens 1981; Stephens and Charnov 1982; Stephens et al. 2007); (2) expected utility theory (Caraco 1980; Real 1980a, 1980b). These two methods are highly effective for characterizing and comparing decision-making processes under conditions of uncertainty and have been applied often in the interpretation of behavior in experimental settings. However, they are not easily applicable towards elucidating the sequence of behavioral decisions or dynamic decision-making.

Optimizing behavior under the constraints of uncertainty and risk often depends on the precise sequence of behavioral decisions or dynamic decision-making. Mathematical approaches towards optimizing dynamic decision-making include dynamic programming (DP) (Bellman 1957) and stochastic optimal control (Pontryagin et al. 1962). Since its introduction to behavioral ecology, dynamic programming has been extensively applied to optimality analyses of behavior and life history strategies (Houston and McNamara 1982; Mangel and Clark 1986, 1988; Houston et al. 1988). Moreover, dynamic programming has proved useful

J. Yoshimura (✉) · H. Ito · K. Tainaka
Department of Systems Engineering, Shizuoka University,
3-5-1 Johoku, Naka-ku, Hamamatsu 432-8561, Japan
e-mail: jin@sys.eng.shizuoka.ac.jp

J. Yoshimura
Department of Environmental and Forest Biology,
College of Environmental Science and Forestry, State University
of New York, Syracuse, NY 13210, USA

J. Yoshimura
Marine Biosystems Research Center, Chiba University,
Uchiura, Kamogawa, Chiba 299-5502, Japan

D. G. Miller III
Department of Biological Sciences, California State University,
Chico, CA 95929, USA

in evaluating the lifetime reproductive success (Darwinian fitness) in the development of individual-based models (Judson 1994).

In contrast, stochastic control theory has only rarely been applied to behavioral and evolutionary ecology (see e.g., Katz 1974; Oaten 1977; Oster and Wilson 1978). This is mainly because of the difficulty and intractability of Pontryagin's maximum principle, the core principle of stochastic control theory (Stephens and Krebs 1986; Mangel and Clark 1988). Further, the applicability of this approach to behavioral biology is rather limited owing to its complexity and the numerous, possibly unrealistic, assumptions inherent in it.

Apart from some additional assumptions of continuous variables (Mangel and Clark 1988), stochastic control theory is mathematically equivalent to dynamic programming, but the latter is more powerful and flexible when applied to dynamic decision-making in behavioral and life history analyses (Mangel and Clark 1986; Houston et al. 1988). The essence of dynamic programming is Bellman's principle of optimality (Bellman 1957, 1961), which states that in a multistage decision process, the series (sequence) of optimal choices always consists of the optimal choice at each time step.

In spite of the wide applicability and usefulness of dynamic programming in behavioral and evolutionary ecology, there is a weak point in this method: the difficulty in the interpretation of numerical results. In dynamic programming, the output is always in the form of numerical tables summarizing dynamic changes of animal behavioral states: characterizing optimal behavior is not straightforward. Therefore, great care must be exercised in the interpretation of numerical values representing optimal behavioral sequences. In some complicated cases, such interpretation is impossible or nearly so, because of complex model settings. In spite of its limitations, dynamic programming has remained the most useful approach for exploring dynamic problems in behavioral and evolutionary ecology, since stochastic control theory is much more limited in its applicability (Mangel and Clark 1988).

Here we develop an analytical model of dynamic decision-making under uncertainty and risks by applying the mathematical analyses (techniques) developed in prior studies of geometric mean fitness (Yoshimura and Clark 1991), within the framework of dynamic utility maximization (Strotz 1955). We argue that this is the first method to characterize properties of fitness in a dynamic sense, in which the fitness function (optimization criterion) changes dynamically in time. In this paper, the mathematical technique developed for geometric mean fitness in stochastic environments (Yoshimura and Clark 1991) is applied to the optimization of sequential decisions of an individual animal. We should note here that the

mathematics we develop for behavioral decision-making is at the individual level within a single generation, and highly distinct from that of geometric mean fitness approximating population growth over many generations. For example, the risk of predation of an individual within one generation is different from that of stochastic variation over many generations (Ito et al. 2013).

We consider the optimization of daily sequential decisions from birth to death of an individual and evaluate its lifetime reproductive success. We consider the optimality of a stochastic decision process by applying Bellman's principle of optimality used in dynamic programming (Bellman 1957, 1961). We solve the optimal decision sequence analytically. The result shows that the fitness function (optimal decision criterion) is dependent on current animal states. Finally we discuss the implications of the result, especially for game theory.

Theory

Suppose that a juvenile animal grows every day to adulthood. Body size, w_t , is the non-negative state variable of the animal (decision-maker) at time t for $t = 0, \dots, T$, where $T \gg 0$ is a finite end time (until death). We may regard body size as a proxy for the energy state or fat reserves of an animal. Note that T corresponds to the time of reproduction in semelparous animals, e.g., anadromous salmon. For simplicity, we assume that the final body size, w_T , represents the potential fitness (lifetime reproductive success of an individual animal).

Let r_t (≥ 0) denote the multiplicative growth rate of body size at time t , such that $w_{t+1} = r_t w_t$. Then the body size, w_t , is expressed:

$$w_t = w_0 r_0 r_1 r_2 \dots r_{t-1} = w_0 \prod_{j=0}^{t-1} r_j \quad (1)$$

Note that, at each time t , the decision-maker chooses an option that results in a growth rate r_t , where r_t is given by the probability distribution of growth rates associated with the option. Then the growth rate r_t ($t = 0, \dots, T$) is a stochastic process, and the decision-maker can optimize this stochastic process by choosing an option at each time point. We here evaluate the optimality of sequential decisions over $t = 0, \dots, (T-1)$ (decision sequence of totalling T decisions). Then the geometric mean $G(r)$ of the body size growth function of an individual in Eq. (1) becomes:

$$G(r) = \prod_{j=0}^{t-1} r_j^{(1/t)} \quad (2)$$

The above Eqs. (1, 2) are similar to those of geometric mean fitness at the population level (Yoshimura and Clark

1991), but differ by incorporating multiple decisions of an individual over its lifetime (at the individual level).

We assume that the multiplicative growth factors r_t ($t = 0, 1, 2, \dots, T$) of the body size of an animal are independent and identically-distributed random variables (*i.i.d.r.vs.*). Here Eq. (2) is the average of observed growth factors r_t evaluated over time t . In order to analyze the optimality of decision sequences, we replace r_t with the corresponding probability distribution $p = p(r)$ (i.e., probability $p(r_j) = 1/t$), such that

$$G(r) = \prod r^{p(r)} \quad (3)$$

Note this equation includes a probability distribution which can be interpreted as an ensemble average estimated by a large number of trials with many individuals. Now the body size w_t at time t (Eq. (1)) is expressed:

$$w_t = w_0 \left(\prod r^{p(r)} \right)^t = w_0 G(r)^t \quad (4)$$

Then we consider the optimization of the body size of an individual until time t , such that

$$\text{Maximize : } w_t \quad (5)$$

for all the decisions over $j = 0, \dots, (t - 1)$. In this paper we call this estimate of reproductive success the potential fitness (of an individual animal). Maximizing w_T (i.e., $t = T$) yields the optimization of body size over an individual's lifetime. This framework is identical to the principle of optimality in dynamic programming (Bellman 1957, 1961). From Eq. (4), maximizing the potential fitness (Eq. 5) becomes simply: $\max \{w_t\} = w_0 \{\max G(r)\}^t$. Then the maximization of the geometric mean potential fitness $G(r)$ is achieved by maximizing $\log\{G(r)\}$, since $\log(r)$ is a monotone increasing function. Thus maximizing potential fitness w_t (Eq. 5) is equivalent to maximizing $\log\{G(r)\}$, such that

$$\begin{aligned} \text{Maximize : } \log G(r) &= \sum_{j=0}^{t-1} \log r_j = \sum_r p(r) \log r \\ &= E\{\log r\}, \end{aligned} \quad (6)$$

where E now refers to the commonly used arithmetic mean, in this case of $\log r$. This expression ties in nicely with utility theory (von Neumann and Morgenstern 1947). Namely we simply define a utility function $u(r)$:

$$u(r) = \log(r) \quad (7)$$

and then express our evolutionary hypothesis as ‘natural selection maximizes expected utility $E(u(r))$ of growth rate r .’ Since $\log r$ is universally concave downwards, we reach the immediate prediction that the utility maximization under conditions of uncertainty is, in principle, *risk averse* (Mangel and Clark 1988, Yoshimura and Clark 1991).

Finally the growth rate r of body size can be replaced by weight gain g and current body size w at any time. Since $r_t = w_{t+1}/w_t$ and $w_{t+1} = (\text{gain at time } t) + w_t$, we get

$$r = \frac{g + w}{w} \quad (8)$$

We then define the dynamic utility surface $u(g; w)$ of gain g (decision variable) given current body size w (state variable), such that:

$$u(g; w) = \log\left(\frac{g + w}{w}\right) \quad (9)$$

and the principle of dynamic utility (DU) optimization is to maximize the expected utility (potential fitness) $E(u(g; w))$ of weight gain g , given current body size w (Fig. 1). Thus, given a current state w , the static solution of dynamic utility becomes a simple form of the expected utility theory with a logarithmic utility function of gain g given w . This result demonstrates that the potential fitness function itself (optimization criterion or utility function) depends on the current state w of animals, i.e., $u = u(g; w)$.

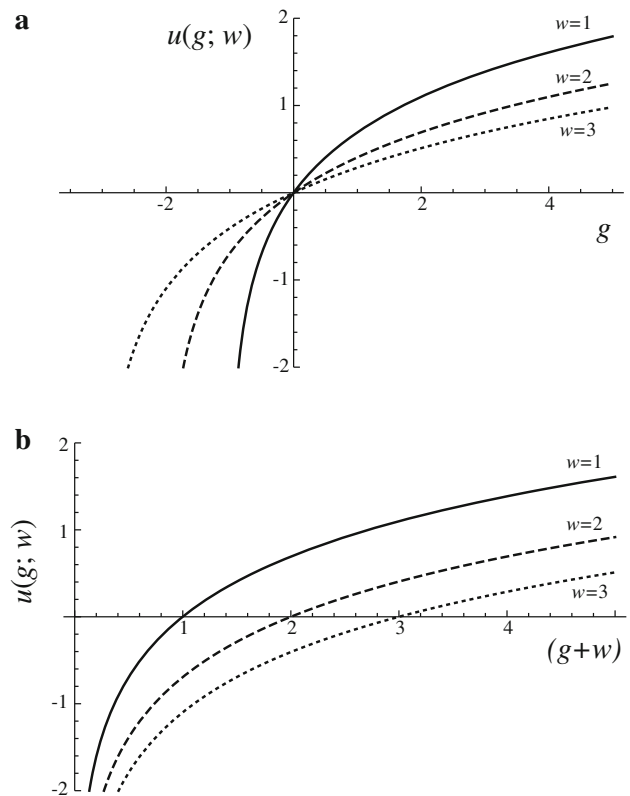


Fig. 1 Dynamic utility function, $u = u(g; w)$. **a**, u plotted against g **b**, u plotted against $(g + w)$. Body sizes are $w = 1$ (solid line), 2 (dashed line), 3 (dotted line)

Discussion

Previously, fitness has been considered a universal function of a single variable w (body size, food, energy or fat reserves), e.g., the utility function $u = u(w)$ (Caraco 1980; Real 1980a, b). However, it has been pointed out in human economic behavior that the shape of the utility function should depend on the current wealth state (Friedman and Savage 1948; Markowitz 1952). Our analysis demonstrates that the shape of the fitness function (potential fitness of an individual) also depends on two variables: one state variable (e.g., current body size) and one decision variable (current gain), i.e., $u = u(g; w)$ (Eq. 9, Fig. 1a). This indicates that potential fitness functions can vary among individual animals as long as their current states (e.g., body size) vary as well (Fig. 1b).

The current result does not agree with the properties of the utility function in a payoff matrix under game theory (e.g., Maynard Smith 1989). For example, consider the payoff matrix of the famous hawk-dove game (Fig. 2). When a dove plays with a dove, they divide the fitness reward, $V/2$. Suppose the body size of one dove (dove 1) is ten times that of the other dove (dove 2), i.e., $w_1 = 10w_2$. Then, the equal fitness value ($V/2$) means that the gain of dove 1 is ten times that of dove 2, i.e., $g_1 = 10g_2$. This implies that the larger one will receive proportionately more food than the smaller one. For example, if there is 110 g of food in total, dove 1 gets 100 g, while dove 2 gets 10 g. This contrasts with the peaceful (equal) division of rewards (55 g each) among doves unless all doves have an equal body size. However, once the game is played among hawks and doves, the body sizes of the players become distinct. The weight w of the dove against a dove becomes $w + V/2$, whereas that of the dove against a hawk remains w . In similar fashion, the hawk body size varies as well. The hawk versus hawk becomes $w + (V - C)/2$, while the hawk versus dove becomes $w + V$. Thus, once the game has begun, the current body sizes of players become highly variable. In the payoff matrix that is derived from the traditional utility theory, the current w does not affect the utility function from the third axiom (so called independence axiom) of expected utility theory (von Neumann and Morgenstern 1947). Thus our result shows that the third axiom is not valid for dynamic behavior in principle. We thus conclude that the expected utility theory is a static model.

In reality, for an animal player with a full stomach, the reward V has close to zero utility, but for a starved animal, it has a considerably higher utility, irrespective of the players being hawk or dove. It follows that the reward v is the actual gain (food or energy gain in animals and money in humans), instead of its utility. This means that the reward is the same for all players in the payoff matrix, but

	Hawk	Dove
Hawk	$(V-C)/2$	V
Dove	0	$V/2$

Fig. 2 Typical payoff matrix of a hawk-dove game. *Dynamic utility* indicates that each payoff should not be a utility, but an absolute gain

the utility of a player depends on the current state (body size or energy reserve) of the player.

Dynamic utility is useful in characterizing behavioral dynamics because its outcome is analytical. Thus it overcomes the limitations of dynamic programming generating only numerical outputs. The analytical solution of dynamic utility derives from the assumption of the multiplicative growth rate r as a stochastic process. However, in analyses of dynamic utility we cannot apply individual environmental parameters used by dynamic programming (or the equivalent stochastic control).

Our analysis proves that the dynamic optimality (optimization criterion) always depends on current states of animals. This indicates that the mean-variance methods and expected utility theory are static models referring to a decision of an individual at one time point (Mangel and Clark 1988; Houston et al. 2011). Dynamic utility is a new addition to the two existing dynamic optimization methods for analyzing animal behavior: dynamic programming and stochastic control.

Most empirical studies of foraging suggest animals exhibit risk-averse behaviors (Caraco and Chasin 1984; Yoshimura and Shields 1987; Ito et al. 2013). Note that the derived dynamic utility function (logarithm of gains) implies a diminishing return of gains. This means that optimal behavioral decisions are universally risk-averse. However, risk-prone behavior has also been observed in some empirical studies in a variety of animal species, including insects (Moses and Sih 1998), fish (Sih 1994), squirrels (Bowers and Breland 1994), chimpanzees (Gilby and Wrangham 2007) and humans (Coddington et al. 2011). The theory of dynamic utility is an important and basic principle of dynamic decision-making and merits further theoretical and empirical contributions.

Acknowledgments This work was supported by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan to J. Y. (nos.22370010 and nos. 22255004) and K. T. (nos.20500204).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Bellman RE (1957) Dynamic programming. Princeton University Press, Princeton
- Bellman RE (1961) Adaptive control processes: a guided tour. Princeton University Press, Princeton
- Bowers MA, Breland B (1994) Foraging of gray squirrels on an urban-rural gradient: use of the GUD to assess anthropogenic impact. *Ecol Appl* 6:1135–1142
- Caraco T (1980) On foraging time allocation in a stochastic environment. *Ecology* 61:119–128
- Caraco T, Chasin M (1984) Foraging preferences: response to reward skew. *Anim Behav* 32:76–85
- Codding BF, Bliege Bird R, Bird DW (2011) Provisioning offspring and others: risk-energy trade-offs and gender differences in hunter-gatherer foraging strategies. *Proc R Soc B* 278: 2502–2509
- Friedman M, Savage LJ (1948) The utility analysis of choices involving risk. *J Polit Econ* 56:279–304
- Gilby RC, Wrangham RW (2007) Risk-prone hunting by chimpanzees (*Pan troglodytes schweinfurthii*) increases during periods of high diet quality. *Behav Ecol Sociobiol* 61:1771–1779
- Houston AI, McNamara J (1982) A sequential approach to risk-taking. *Anim Behav* 13:1260–1261
- Houston A, Clark C, McNamara J, Mangel M (1988) Dynamic models in behavioural and evolutionary ecology. *Nature* 332:29–34
- Houston AI, Higginson AD, McNamara JM (2011) Optimal foraging for multiple 15 nutrients in an unpredictable environment. *Ecol Lett* 14:1101–1107
- Ito H, Uehara T, Morita S, Tainaka K, Yoshimura J (2013) Foraging behavior in stochastic environments. *J Ethol* 31:23–28. doi: [10.1007/s10164-012-0344-y](https://doi.org/10.1007/s10164-012-0344-y)
- Judson OP (1994) The rise of the individual-based model in ecology. *Trends Ecol Evol* 9:9–14
- Katz PL (1974) A long-term approach to foraging optimization. *Am Nat* 108:758–782
- Mangel M, Clark CW (1986) Towards a unified foraging theory. *Ecology* 67:1127–1138
- Mangel M, Clark CW (1988) Dynamic modeling in behavioral ecology. Princeton University Press, Princeton
- Markowitz HM (1952) The utility of wealth. *J Political Economy* 60:151–158
- Maynard Smith J (1989) Evolutionary genetics. Oxford University Press, Oxford
- Moses JL, Sih A (1998) Effects of predation risk and food availability on the activity, habitat use, feeding behavior and mating behavior of a pond water strider, *Gerris marginatus* (Hemiptera). *Ethology* 104:661–669
- Oaten A (1977) Optimal foraging in patches: a case for stochasticity. *Theor Popul Biol* 12:263–285
- Oster GF, Wilson EO (1978) Caste and ecology in the social insects. Princeton University Press, Princeton
- Pontryagin LS, Boltyanskii VG, Gamkrelidze RV, Mishchenko EF (1962) The mathematical theory of optimal processes, Wiley-Interscience, (translated from the Russian)
- Real LA (1980a) Fitness, uncertainty, and the role of diversification in evolution and behavior. *Am Nat* 115:623–638
- Real LA (1980b) On uncertainty and the law of diminishing returns in evolution and behavior. In: Staddon JER (ed) Limits to action: the allocation of Individual behavior. Academic Press, New York, pp 37–64
- Real LA, Caraco T (1986) Risk and foraging in stochastic environments. *Ann Rev Ecol Syst* 17:371–390
- Sih A (1994) Predation risk and the evolutionary ecology of reproductive behaviour. *J Fish Biol* 45A:111–130
- Stephens DW (1981) The logic of risk-sensitive foraging preferences. *Anim Behav* 29:628–629
- Stephens D, Charnov EL (1982) Optimal foraging: some simple stochastic models. *Behav Ecol Sociobiol* 10:251–263
- Stephens DW, Krebs JR (1986) Foraging theory. Princeton University Press, Princeton
- Stephens DW, Brown JS, Ydenberg RC (eds) (2007) Foraging: behavior and ecology. Chicago University Press, Chicago
- Strotz RH (1955) Myopia and inconsistency in dynamic utility maximization. *Rev Econ Studies* 23:165–180
- von Neumann J, Morgenstern O (1947) The theory of games and economic behavior, 2nd edn. Princeton University Press, Princeton
- Yoshimura J, Clark CW (1991) Individual adaptations in stochastic environments. *Evol Ecol* 5: 173–192, 430 (Corrigenda)
- Yoshimura J, Shields WM (1987) Probabilistic optimization of phenotype distributions: a general solution for the effects of uncertainty on natural selection? *Evol Ecol* 1:125–138